Krebs, W.K., McCarley, J.M., Kozek, T., Miller, G.M., Sinai, M.S., and Werblin, F.S. (1999). An evaluation of a sensor fusion system to improve drivers' nighttime detection of road hazards. Proceedings of the 43rd Annual Meeting Human Factors and Ergonomics Society, 43, 1333-1337.

AN EVALUATION OF A SENSOR FUSION SYSTEM TO IMPROVE DRIVERS' NIGHTTIME DETECTION OF ROAD HAZARDS

Krebs[1], WK, McCarley[1], JS, Kozek[2], T, Miller[3], GM, Sinai[1], MJ, and Werblin[2,4], FS

[1]Naval Postgraduate School, Department of Operations Research, Monterey, CA; [2]TeraOps Corporation, Berkeley, CA; [3]PVP Advanced EO Systems Inc, Orange, CA; [4]University of California, Berkeley, Molecular Cell Biology Department, CA

Objective: To determine whether combination of two spectral bands within a single sensor-fused image can improve drivers' detection of road hazards. **Methods**: Images were collected with visible and short wave infrared sensors, and were combined by an image fusion algorithm derived from a computational model of human retinal processing (Werblin et al., 1997). Test stimuli were single-band and fused images of a nighttime scene, collected with sensors mounted atop a vehicle and facing down a stretch of road into an opposing vehicle's headlights. The intensity of the opposing headlights was varied to produce images of low and of high glare. The subjects' task was to detect the presence of a pedestrian within each image. **Results:** Sensor-fused imagery reliably produced performance better than or equivalent to that produced by either format of single-band imagery. **Conclusions:** Sensor fusion may provide an effective method of facilitating the detection of road hazards under low visibility conditions.

INTRODUCTION

General Motors Corporation will integrate an infrared (IR) sensor suite as an optional feature on next year's Cadillac DeVille, with the objective of improving drivers' nighttime visibility by three to five times under low beam illumination (General Motors Corporation, 1998). An uncooled long-wave IR sensor will be located in the vehicle's front grill, and its output image will be displayed on an HUD (11^0 horizontal and 4^0 vertical) in the lower part of the windscreen just above the steering wheel. The IR sensor, responsive to thermal energy, will be meant to complement the headlights during night driving, and might be particularly useful for detecting objects that are not illuminated by the sensor vehicle's headlights, or that are masked by glare of other vehicles' headlights. To utilize the IR display, however, drivers will be required to alternate attention and gaze between the front-viewed scene and the HUD, and to integrate information from displays differing in size, aspect ratio, luminance and spatial detail.

An alternative method of improving nighttime driver's performance might be through presentation of visible and IR information within a unitary, *sensor-fused* image. Such a fused image could allow observers to simultaneously view the output of multiple sensors, obviating the need for shifts of attention or fixation between multiple displays. Additionally, sensor-fusion could allow electro-optically-sensed data to be rendered chromatically. Because visible and IR sensors differ in spectral sensitivity, the images they produce generally have grossly similar spatial structure, but differ in pixel by pixel intensity. This contrast between the output of multiple sensors can be taken as the basis of a color-mapping scheme, much as differences in the output of retinal receptors provide basis of human color vision. Such color mapping should improve image contrast and allow drivers to more quickly detect and orient themselves to approaching road

Krebs, W.K., McCarley, J.M., Kozek, T., Miller, G.M., Sinai, M.S., and Werblin, F.S. (1999). An evaluation of a sensor fusion system to improve drivers' nighttime detection of road hazards. Proceedings of the 43rd Annual Meeting Human Factors and Ergonomics Society, 43, 1333-1337.

hazards. Currently, the United States Army intends to install a color night vision system into some of their ground vehicles.

The purpose of the present study was to determine, first, whether sensor-fused images combining the output of visible-light and short-wave IR (SWIR) sensors can improve subjects' ability to detect road hazards (here, a pedestrian) within a nighttime scene, and second, whether chromatically fused images can support better performance than can achromatically fused images. Observers performed a psychophysical task that required them to detect a target object within naturalistic nighttime scenes rendered in single-band, achromatic sensor-fused dual-band, and chromatic sensor-fused dual-band formats. Images were degraded with either low or high level of glare. It was hypothesized that target detection within fused images would be faster and more accurate than within single band images, and that color-fused imagery would support better psychophysical performance than would achromatic fused imagery. Further, it was hypothesized that the benefits of sensor fusion would be greater under conditions of high glare. These results may assist automobile manufactures' in assessing the utility of color fusion systems.

METHODS

Observers: Eleven military observers with normal or corrected to normal vision volunteered for this experiment. Informed consent was obtained from all observers.

Stimuli: Nighttime driving images were collected using vertically mounted visible and SWIR cameras secured to the top of a sport utility vehicle. The two cameras were spatially registered with a 40^0 field-of-view. The visible camera (Pulnix TM-745I) had a 762 by 498 pixel element array with a 2/3" format, and had a spectral sensitivity between 400nm-700nm with a peak sensitivity of 500nm. The SWIR camera (Pulnix TM-540) had a 510 by 492 pixel element array with a 2/3" format, and had spectral sensitivity between 700nm-1250nm with peak sensitivity of 1000nm. Images depicted a length of road, roughly parallel to the sensors' line of sight, with an opposing vehicle stationed at a distance of 45.72 m and facing directly into the sensors' field of view. The sensor vehicle's headlights were set to low beam and the opposing vehicle lights alternated between low beam and high beam illumination. The target object, a pedestrian person crossing the road perpendicular to the test vehicle, was positioned at a distance of either 30.5 meters or 38.1 meters from the test vehicle.

Twelve target-present images were digitized, six containing a near target and six containing a distant target. The target's lateral position varied between images. Scenes with no target present served as noise trials. Multiple target-absent scenes were used to avoid the possibility that observers might base their responses on differences in ambient background noise between target-present and target-absent scenes.

Images were presented in six different formats: two single-band formats (visible and SWIR), two chromatic sensor-fused dual-band formats, and two achromatic sensor-fused dual-band formats (refer to figure 1). The sensor-fusion method used (Werblin, Roska, Chua, Jacobs, Kozek and Zarandy, 1997) employs a color-rendering algorithm that combined the original visible and SWIR images via processing similar to that represented in a computational model of the human retina which emphasizes the spatial and temporal diffusion of local responses of different classes of neurons in the human retina (e.g., the time-course of retinal center/surround spatial antagonism; Werblin, Roska, and Chua, 1995; Jacobs, Roska, and Werblin, 1996). The fusion processing is based on computations that rely on only nearest-neighborhood interactions of pixels, suitable for realization on a Cellular Neural Network (CNN) processor (i.e., a network of analog processors, locally connected, with nonlinear dynamics; see Chua and Roska, 1993). The algorithm used to create gray-scale fused images employs a sequence of spatial-temporal CNN operators that combine information from input channels in a locally adaptive manner. The resulting images were subsequently pseudo-colored by a procedure that assigns hues to approximate natural appearance of objects based on target and sensor

Krebs, W.K., McCarley, J.M., Kozek, T., Miller, G.M., Sinai, M.S., and Werblin, F.S. (1999). An evaluation of a sensor fusion system to improve drivers' nighttime detection of road hazards. Proceedings of the 43rd Annual Meeting Human Factors and Ergonomics Society, 43, 1333-1337.

characteristics. Two distinct color mappings were employed to create images of two color-fused formats, color1 and color2. Achromatic dual-band formats gray1 and gray2 were created by converting chromatic dual-band images of formats color1 and color2, respectively, to grayscale. Achromatic dual-band images served as controls for chromatic dual-band images, and were employed to allow the effects of fused spatial information to be distinguished from the effects of false-color within the fused imagery.

Procedure: Observers were asked to determine whether a target was present or absent within each stimulus image, and to indicate their responses manually by pressing "1" (target present) or "2" (target absent) on the computer keyboard's numeric keypad. At the onset of each trial, the subject fixated on a cross hair located in the center of the screen. The fixation cross remained visible for 500 msec, and was followed immediately by presentation of the experimental stimulus, which remained visible until subject's response. The next trial began approximately 1 second after the preceding trial's end. Reaction time and accuracy were measured for each trial. To prevent observers from learning to make responses based on idiosyncratic features within individual images, no feedback was provided. An experimental session consisted of two blocks of 216 trials each. Each block contained 144 target-present trials, with each of the twelve high-glare and low-glare target-present images presented in each of the six formats, and 72 target-absent trials, six of high-glare and six of low-glare in each of the six formats.

RESULTS

Reaction times (RTs) for incorrect responses were discarded. Initial examination of data revealed that RTs were reliably faster for target-present than for target-absent responses, and were faster with low-glare than with high-glare imagery. Analysis revealed no reliable main effects of image format on RTs, however, and comparison of RTs and accuracy rates produced evidence of a speed accuracy tradeoff for only one condition, the high-glare SWIR imagery. Further analysis, qualified by evidence of this speed accuracy tradeoff for one condition, therefore focused exclusively on observers' accuracy rates.

Accuracy data for each combination of image format and level of glare was used to calculate A', a measure of sensitivity similar to d' (Macmillan and Creelman, 1991). A' was used to calculate the area under the one-point ROC. In this study, each subject contributed one point on the ROC curve. Figure 2 illustrates the mean A' for each image format for both high and low glare conditions.

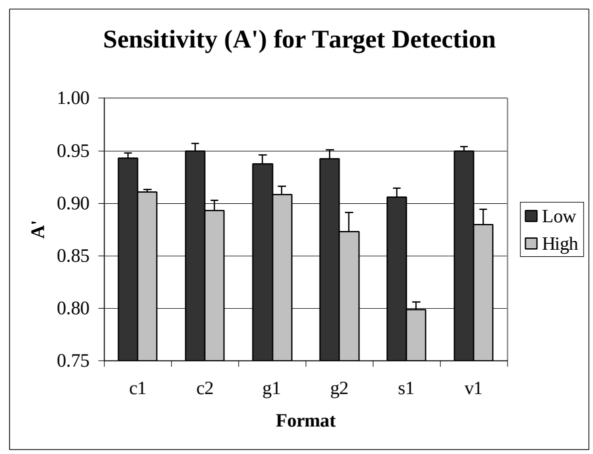

Figure 2. Subjects' mean sensitivity for the six different display formats at two levels of glare. Short-wave infrared (s1) was significantly worse compared to the other five formats for both low and high glare illumination conditions. Collapsing color fusion color1 (c1) and color2 (c2) was not significantly different from grayscale fusion gray1 (g1) and gray2 (g2), however color and grayscale fusion was significantly better than both single bands in the high glare condition.

A 6 x 2 within-subjects ANOVA with image format and level of glare as factors revealed a significant main effect for display format $F(5, 50) = 18.95$, $p < .0001$). Pairwise post-hoc tests (protected t-test HSD, df=50, $\alpha = .05$) for the high beam condition showed that performance with dual-band imagery of format color1 or gray1 was reliably better than performance with either single-band visible imagery or with imagery of dual-band format gray2. Performance did not differ reliably between fused imagery of formats gray1 and color1, however, suggesting that this fusion algorithm facilitated visual performance by improving the spatial content of the stimulus images, and not by enhancing their contrast through the addition of chromaticity. Post-hoc t-tests also indicated that observers were reliably less accurate with SWIR imagery than with imagery of any other format. Since performance with high-glare SWIR imagery gave evidence of a speed accuracy tradeoff, unfortunately, these differences are ambiguous. Post-hoc tests (HSD, df=50, $\alpha = .05$) for the low beam condition showed that performance was reliably worse with single-band SWIR imagery than imagery of any other format, suggesting that fusion of single-band SWIR and single-band visible imagery effectively retained information of the visible displays. No further reliable differences emerged between imagery of various formats under low levels of glare.

CONCLUSION

Under conditions of low-glare, sensor fused imagery produced performance equal to or better than that produced by either single-band format, while under conditions of high glare, sensor-fused imagery produced performance better than that produced by either form of single-band imagery. In neither case did color rendering of sensor-fused imagery improve performance relative to that which obtained with achromatically rendered sensor-fused images. Results suggest that sensor-fusion can effectively combine, and perhaps enhance, spatial information provided by multiple single-band sensors, and that fusion may offer a method of facilitating the detection of road hazards under conditions of poor visibility.

ACKNOWLEDGEMENTS

A special thanks to Mr. George Jaksha, Physics Department at the Naval Postgraduate School, for building the sensor platform in a timely manner. Sponsored by DARPA's Integrated Imaging Sensors Program, Mr. Ray Balcerak, program manager. Also, sponsored by Office of Naval Research contract #N0001497WR30091, #N0001498WR30112.

REFERENCES

Chua, L.O. & Roska, T. (1993). The CNN paradigm. IEEE Transactions on Circuits and Systems, 40, 147-156.

General Motors Corporation, Cadillac Division. Technical backgrounder: Cadillac's night vision system. (1998, August 20). Available at: http://www.cadillac.com

Jacobs, A., Roska, T. and Werblin, F. (1996). Methods for constructing physiologically motivated neuromorphic models in CNNs. International Journal of Circuit Theory and Applications, 24, 315-339.

Macmillan, N.A. & Creelman, C.D. (1991). Detection theory: A user's guide. Cambridge: Cambridge University Press

U.S. Department of Transportation, National Highway Traffic Safety Administration. (1996). Motor vehicle crash data from FARS and GES. Available at http://www.nhtsa.dot.gov

Werblin, F.S. Roska, T. and Chua, L. (1995). The analogic cellular neural network as a bionic eye. International Journal of Circuit Theory and Applications, 23, 541-569.

Werblin, F.S., Roska, T. Chua, L., Jacobs, A., Kozek, T., and Zarandy, A. (1997). Transfer of retinal technology to applications beyond biological vision. Investigative Ophthalmology and Visual Science, (Suppl) 38, 479.

Krebs, W.K., McCarley, J.M., Kozek, T., Miller, G.M., Sinai, M.S., and Werblin, F.S. (1999). An evaluation of a sensor fusion system to improve drivers' nighttime detection of road hazards. Proceedings of the 43rd Annual Meeting Human Factors and Ergonomics Society, 43, 1333-1337.

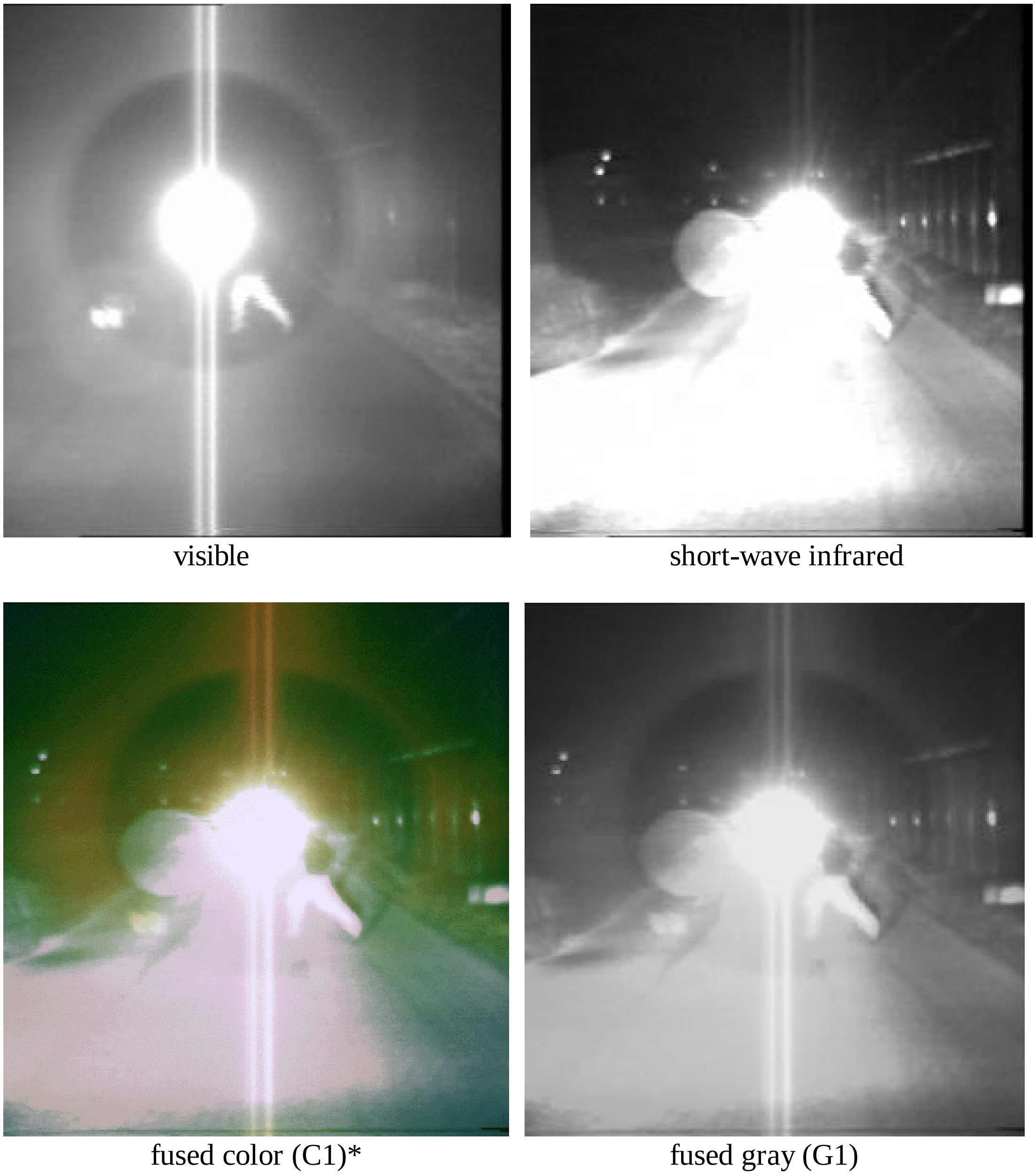

Figure 1. The opposing vehicle lights were set at high beam illumination with a pedestrian located in the center of the road 30.5 meters from the sensor vehicle. Observers were asked to determine whether a pedestrian was present or absent within each stimulus image.

***fused color (C1) condition is printed in black and white.**